A PORTABLE LOW VISION AID BASED ON GPU

R. Ureña, P. Martinez-Cañada, J. M. Gómez-López, C. Morillas and F. Pelayo
Departamento de Arquitectura y Tecnología de los Computadores, CITIC, ETSIIT
Universidad de Granada, C/ Periodista Daniel Saucedo Aranda s/n, Granada, Spain

Keywords: Low vision, Head-mounted display, Real time, Video processing, GPU, CUDA, GPGPU.

Abstract: The purpose of this work is to describe a customizable aid system based on GPU for low vision. The system aims to transform images taken from the patient's environment and tries to convey the best information possible through his visual rest, applying various transformations to the input image and projecting the processed image on a head-mounted-display, HMD. The system easily enables implementing and testing different kinds of vision enhancements adapted to the pathologies of each low vision affected, his particular visual field, and the evolution of his disease. We have implemented several types of visual enhancements based on extracting an overlaying edges, image filtering, and contrast enhancement. We have developed a complete image processing library for GPUs compatible with CUDA in order the system can perform real time processing employing a light-weight netbook with an integrated GPU NVIDIA ION2. We briefly summarize here their computational cost (in terms of processed frames per second) for three different NVIDIA GPUs.

1 INTRODUCTION

There are thought to be 38 million people suffering from blindness worldwide, and this number is expected to double over the next 25 years. Additionally, there are 110 million people who have severely impaired vision. (Fosters A. et Al, 2005)

Low Vision (LV) is the term commonly used to describe partial sight, or sight which is not fully correctable with conventional methods such as glasses or refractive surgery (Peláez-Coca et al., 2009).

The low vision pathologies can be divided mainly into two categories; those that predominantly suffer from a loss of visual acuity due to macular degenerations, and those that suffer from a reduction in the overall visual field such as Retinitis Pigmentosa. In many countries, there is an increasing prevalence of diabetic retinopathy and an ageing population with 1 in 3 over the age of 75 being affected with some form of ageing macular degeneration (Al-Atabany et al., 2009). The affected who suffer from loss of visual acuity lose their foveal vision and therefore they can walk easy and avoid objects but they struggle to read or to watch TV. On the other hand, those who suffer from a reduction of their visual field, such as Retinitis Pigmentosa, can perform properly static tasks which require a relatively reduced visual field. However their ability to walk and avoid obstacles is very limited. Moreover they experience a progressive loss of contrast sensitivity and therefore they have many problems to manage themselves in low illumination environments or, in general, in environments where the illumination is not controlled.

There are several LV aids which try to improve the visual capabilities taking advantage of residual vision. Some of these devices employ an opaque and immersive HMD to project the enhanced images. For example the System LVES (Massof et al., 1992), and the system JORDY by Enhanced Vision. Also it has been developed a portable aid systems (Peli. E, 2001; Peláez-Coca et al., 2009) based on see-through displays which overlap the edges of the whole scene over the patient's useful visual field. These systems are specially oriented to Retinitis Pigmentosa affected people and use DSP devices and/or FPGA for real-time processing.

The aid systems described above perform transformations of the input image, amplifying it in size, intensity or contrast. These transformations are mainly based on digital zooming and edge overlaying. The systems based on the magnification of the image are very useful in static and controlled

Ureña R., Martinez-Cañada P., Gómez-López J., Morillas C. and Pelayo F. (2011).
A PORTABLE LOW VISION AID BASED ON GPU.
In *Proceedings of the 1st International Conference on Pervasive and Embedded Computing and Communication Systems*, pages 201-206
DOI: 10.5220/0003366702010206

illumination environments, such as watching television or reading but they have little use in mobile environments, since they reduce the visual field and present unrealistic images which prevent the user from getting a real insight into the distance at which obstacles are.

Most of the LV pathologies are characterized by a slow progression with residual vision deteriorating gradually with time; therefore the patients have requirements that change as the disease advances. Moreover the LV diseases affect unevenly to different areas of visual field, thus a non-uniform processing adapted to the affected needs and visual field may be useful.

The systems mentioned above do not enable totally customize the processing to the visual needs and disease progression.

In this context the main contribution of the present system is a new platform with allows implementing and testing different kinds of image enhancements adapted to the visual needs of each affected, to his visual field, and to the evolution of his disease. So as to customize the enhancements the system has a graphical user interface. Moreover we have developed different kinds of image enhancements which improve the image contrast even in low light environments where low vision affected experiment several difficulties. The designed system achieves real time image processing (above 25 frames per second video-rate) using a last generation Graphic Processor Unit (GPU) integrated in a light weight netbook.

Even though embedded solutions based on DSPs and/or FPGA may provide speed performance, modern GPUs integrated in small size portable computers can also provide the minimum latency and frame rate required as they have multiple scalar processors. The main advantage of GPU-based systems is that they are easier and faster to customize to the needs each visual impaired than other implementations. It also provides facilities for rapid development and testing of new image enhancements.

2 SYSTEM SPECIFICATIONS

The proposed system can be viewed as a SW/HW platform for low vision support, which aims to easily implement and test different types of visual correctors tailored to the needs of each affected, and his visual field. Therefore the system aims to transform images taken from the patient's environment and tries to convey the best information possible through his visual rest, applying different transformations to the input image.

The main characteristics are:

(1) **Customizable System:** The system is able to perform a sequence of transformations totally adapted to the visual requirements, and visual field of each low vision affected.

(2) **Portability:** The image processing device needs to be carried by the patient in mobile environments such as walking and similar tasks.

(3) **Real Time Processing:** The system is able to perform different image enhancements in real-time by using a low-power GPU embedded in a light weight netbook.

(4) **Flexibility:** The system can combine several types of visual enhancements including digital zooming, spatial filtering, edge extraction and tone-mapping and works properly in non uniform illumination environments.

2.1 Architecture

The developed platform runs over a netbook ASUS EEPC 1201 PN. It uses the netbook's CPU and a GPU NVIDIA ION2 connected via PCI-express.

In the CPU runs the main application, and is where the user can define the processing to be performed according to the visual needs of each LV using a graphical user interface (UI). The UI is based in the system RETINER (Morillas et al., 2007) and a platform for speeding up non-uniform image processing (Ureña et al., 2010). The application performs algebraic optimizations based on the convolution properties to simplify filter stages.

After the optimization we can make out what tasks are to run on the GPU and on the CPU. The tasks performed by the CPU are invoked directly by the application, whereas in the case of the GPU using MEX (NVIDIA Corporation, 2007) modules allows us to both set the type of processing to be performed, and image transfers.

In Figure 1 we can see a diagram that summarizes the functional architecture of the implemented system.

Our system uses GPU to speed up the image processing since current GPUs has a multiprocessor architecture suitable for pixel-wise processing.

Most GPUs, given its size and high power consumption are not suitable for portable applications. However, the GPU used in this system, the NVIDIA ION2, has 16 processors integrated on a platform with low power consumption; which has its own battery with about 4 hours of usage.

Moreover the system takes advantage of the Intel ATOM N450 processor, integrated in the used netbook, which is faster than other processor buit-in FPGAs, for example PowerPC.

The system can receive as input live video captured from a camera, images and videos in AVI format. In all cases the image can be in grayscale or color. The color scheme can be RGB, YCbCr or HSV performing the conversion between these color schemes automatically. The system output is a processed image in grayscale or RGB format, depending on the input image and the particular characteristics of the processing chain. The processed image is projected on a Head-Mounted-Display (HMD). In figure 2 we can see a person using the system. As we can see the system have an USB camera, and a HMD display, both connected to the netbook ASUs EEPC 1201PN.

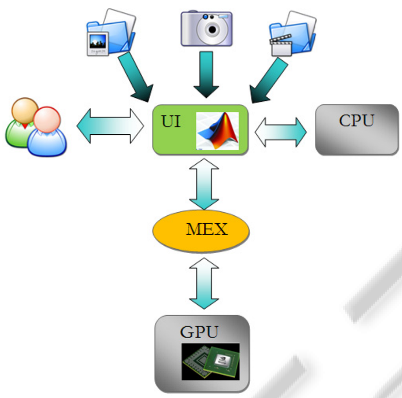


Figure 1: System Architecture.

Figure 2: Example of a person using the system.

2.2 Available Image Enhancements

The various image enhancements that can be applied in this version of the platform are:

Edge Detection: Edge detection and overlapping has been used widely in low vision rehabilitation with patients with central and/or peripheral vision loss (Peli, 2007). It have been assessed the effect of this enhancement on performance and on perceived quality of motion video. The results indicate that adaptive enhancement (individually-tuned using a static image) adds significantly to perceived image quality when viewing motion video.

Contrast-Enhancement: Most of the low vision affected people experiment a noticeable loss of contrast sensitivity, resulting in almost a complete loss of vision in low light environments or in environments where the illumination is not totally controlled (sudden changes in lighting conditions, for example). One of the main objectives of this system is to help the affected precisely in these environments. Therefore we have developed a new method to improve image contrast, based on the conversion of the image to the HSV color space. The system calculates automatically the histogram of the component V to detect if the captured image is too dark, too light or well contrasted. Then equalization of the V channel or of the S channel is done if the image is too dark or too light respectively. Finally the enhanced channels are combined linearly with the original ones using a weighting factor set by the user depending on the desired degree of enhancement. Also we have included a tone-mapping operator (Biswas, et al., 2005).

Figure 3 (a) shows a sunset, in which many of the characteristics of the image have been lost while, in the enhanced images, we can appreciate all the details of the landscape, such as the threes.(See Figure 3 (b) and Figure 3 (c)) Figure 3 (d) shows a man driving, as we can see all the details of the face, such as the ear, have been lost nevertheless we can see clearly the things which are outside the car. In the enhancement images (Figure 3 (e) and (f)) we can appreciate all the details of the face, and also we can see clearly the details of the street. Comparing the Biswas algorithm with the one presented in this article we can see that the former clarifies more the picture distorting in some places the color of the image (see the sky tone in Figure 3 b and the face tone in Figure 3 e) whereas the one presented here even enhanced the colors.

Other Image Enhancements: The system also can perform other image processing tasks like digital zooming, spatial filtering with several types of masks (Gaussian, Difference of Gaussian, Laplacian of Gaussian, Unsharp). The Unsharp mask is of special interest in low vision context since it provides edge and contrast enhancement. Also the system can perform histogram calculation and equalization; these transformations are useful for contrast enhancement, and for automatic thresholding.

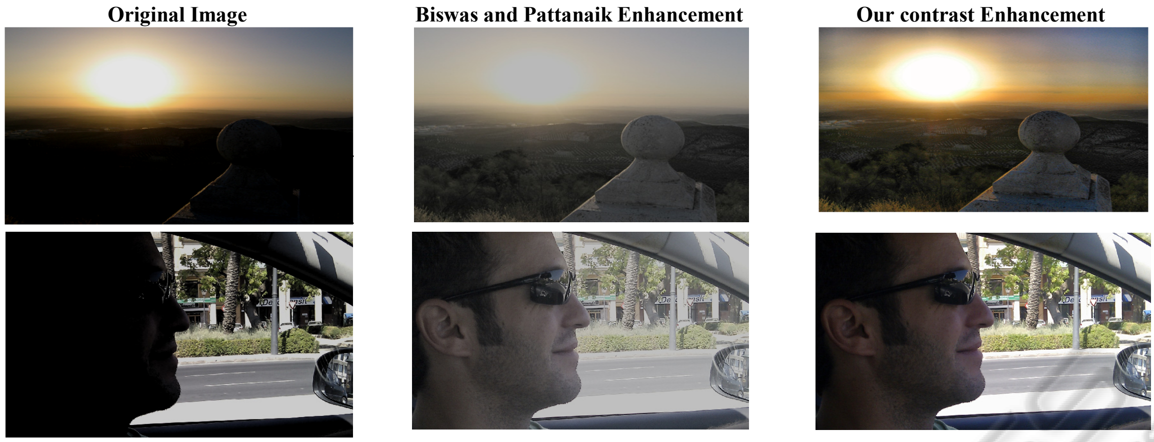
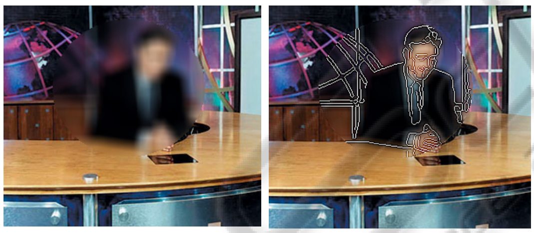


Figure 3: Contrast enhancement examples.

2.3 Non Uniform Processing and Simplification

Many LV pathologies affect unevenly to different regions of the visual field. Therefore, in some cases, it could be useful to perform different kinds of visual enhancements depending on the specific region of the visual field. For example, a Macular Degeneration affected person suffers from a partial loss in his foveal vision, which varies depending on the evolution of the disease, whereas his peripheral vision is undamaged (see Figure 4. a)



(a) (b)

Figure 4: Example of non uniform processing.

Therefore, in certain situations, he may need the central region of his visual field to be enhanced, but it is not necessary to perform any kind of processing in the peripheral region. In figure 4.b we can see an example of non-uniform processing. We have enhanced the edges only in the region where the LV affected presents vision loss.

Each LV affected has different regions of interest (ROI) according to the specific characteristics and the evolution of the disease. Consequently our system has a graphical tool with enables defining different types of ROIs adapted to the visual field of each affected.

All the image enhancements available in the system and explained in section 2.2 can be combined. So as to define a complete processing chain, the user may introduce a text chain which specifies how the different transformations are combined and the ROI to each transformation must be applied.

To combine the transformations the system has three operators which are explained in table 1.

Table 1: Available operators.

Operator	Function
+	Sums the output from the implied transformations.
-	Subtracts the output from the implied transformations.
,	Concatenates transformations.

Once the processing chain is defined, the system performs an algebraic simplification, if necessary, so as to minimize the number of filtering stages. The simplification is based in the convolution properties and enables reducing N consecutives or parallel filtering stages in one stage with a mask resulting from the convolution or the sum /subtraction of the N masks respectively. In order to do all the possible simplifications the system changes the transformations order if possible, taking into account that contrast transformations are not commutable with filters.

2.4 Real Time Processing using GPU

To perform all the image enhancements mentioned above in real time we have developed an extensive library of processing modules for GPU in CUDA (NVIDIA Corporation, 2009).

Our target GPU, the NVIDIA ION2 , consists of two streaming multiprocessors. Each streaming multiprocessor has one instruction unit, eight stream processors (SPs) and one local memory (16KB). Thus it has 16 SPs in total. Eight SPs in a streaming multiprocessor are connected to one instruction unit. This means that the eight SPs execute the same instruction stream on different data (called *thread*). In order to extract the maximum performance of SPs by hiding memory access delay, we need to provide four threads for each SP, which are interleaved on the SP. Therefore, we have to provide at least 32 threads for each streaming multiprocessor.

To optimize the use of the available multiprocessors, the parameters to be determined are the number of threads/block and the shared memory space between the threads of each block.

To accurately size the modules we have used the CUDA Occupancy Calculator tool that shows the occupation of the multiprocessor's cache and its percentage of utilization. The thread block size is chosen in all cases so that multiprocessor occupancy is 100%. The size of the GRID (number of processing blocks to be executed by the kernel) is set dynamically according to the size of the image.

The streaming multiprocessors are connected to large global memory (512MB in ION2), which is the interface between the CPU and the GPU. This DRAM memory is slower than the shared memory, therefore at the beginning of the module all the threads of a block load in the shared memory the fragment of the image that this block needs. Depending on how the data are encoded in the GPU global memory, each thread can load 1 data if we are working with 4-byte data or 4 data if we are working with 1-byte data. The global memory accesses of the GPU for both reading and writing are done so that in one clock cycle all the threads of a warp (K) access to 4K bytes of RAM, ,where K is equal to 32 if we work with CUDA Compute Capability GPUs 1.x.

Before turning to the processing stage all the threads of the processing block have to wait in a barrier to ensure that all of them have loaded its corresponding data. Following the calculation step may be a second stage of synchronization of the block threads before writing to the GPU global memory. The general structure of the GPU processing modules is illustrated in Figure 5.

The interface between the host and the GPU global memory is the bottleneck of the application so each image data is encoded as 1-byte unsigned integer. Therefore to encode a color pixel 3 bytes are used. If more precision is needed (when working with HSV color space for example) a conversion to float is done once the image is stored in the GPU global memory, exploiting the parallelism provided by the GPU.

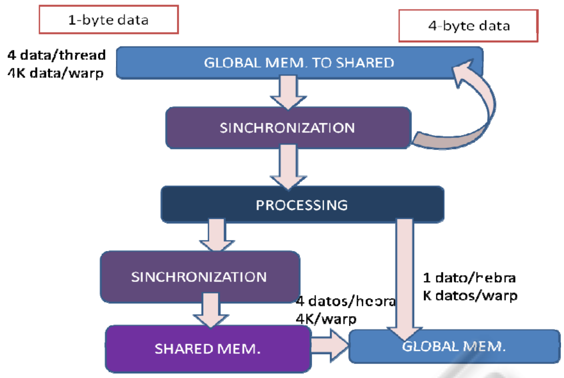


Figure 5: General structure of the GPU modules.

2.4.1 GPU Modules Performance

In this section we present the performance of the GPU modules in terms of frames per second (fps).
For this evaluation we use three different platforms to verify the scalability to the number of multiprocessors available in each GPU:

1. GPU NVIDIA ION2 512 MB DDR3, 2 streaming multiprocessors.
2. GPU NVIDIA GeFORCE 8800GT 512 MB DDR3, 14 streaming multiprocessors.
3. GPU NVIDIA 9200MGS, 256MB DDR3, 1 streaming multiprocessor.

In the measures are not included the image transfer delay from host to GPU global memory and vice versa, which are approximately 2.6 ms for the NVIDIA GeForce 8800, 4.1ms for the NVIDIA GeForce 9200MGS and 12 ms for the NVIDIA ION2 when working with 800x600 RGB images, applying the transformation to the whole image.
As we can see in table 2 in the case of the NVIDIA GPU ION2 all developed modules work in real time, more than 25 frames per second. If we combine several transformations the total processing delay is the result of summing the processing delay of each transformation and the transference delay.

3 DISCUSSION AND FUTURE WORK

We have presented a portable system that enables in such an easy an effective way to combine and test a wide range of visual enhancements of utility for low vision affected that can benefit from on-line real-time image processing.

One of the main advantages of the system is that it can be fully customized to particular user's requirements, such as visual field or the evolution of the pathology, covering a wide range of visual disabilities. In order to adapt the system to the visual field of each LV affected, the platform is able to perform a specific processing to each region of the visual field. Furthermore it can work properly in not controlled or even in low illumination environments since it is able to carry out real-time contrast enhancement algorithms, and it allows the incorporation of other visual enhancements (which might be proposed and tested by others authors).

Table 2: Performance of the GPU modules.

	FPS GPU 9200MGS	FPS GPU 8800GT	FPS GPU ION 2
Filtering (mask size 7x7)	25.71	200	54.26
Histogram equalization	71.68	625	126.9
Edge detection	28.22	370.37	50.4
LUT substitution	221.73	333.33	389.11
RGB to HSV	54.14	312.5	98.14
RGB to YCbCr	91.83	476.19	161.55
Digital zooming	20.96	270.27	37.01
Tone-Mapping Biswas	19.67	229.89	34.65
Contrast enhancement based on HSV	34.6	338.98	60.06

The system can be used in mobile environments such as walking since it is able to perform real time processing in a light weight netbook. In order to achieve real time processing we have developed an image processing library for GPUs compatible with CUDA. The performance of each GPU module in terms of frames per second have been measured for three different GPUs, our target GPU, the NVIDIA ION2, and two more, to show the scalability of the developed modules to the number of multiprocessors available in each GPU.

In order to demonstrate the usefulness of this unique visual aid system we are going to conduct a series of tests with a group of Retinitis Pigmentosa affected. Specifically for this group, the platform is going to be used to enhance image features in low contrast environments where those affected experience several difficulties.

Furthermore we are planning to implement some of the most useful visual enhancements using others embedded devices based on ARM processor, or FPGA, in order to obtain real-time processing in smaller and lighter devices than heretofore employed netbook.

ACKNOWLEDGEMENTS

This work has been supported by the Junta de Andalucía Project P06-TIC-02007, and the Spanish National Grants RECVIS (TIN2008-06893-C03-02) and DINAM-VISION (DPI2007-61683).

REFERENCES

Al-Atabany W., Memon M., Downes S. M., Degenaar P. A., 2010 Designing and testing scene enhancement algorithms for patients with retina degenerative disorders, *BioMedical Engineering OnLine.*

Biswas K., Pattanaik S., 2005 Simple Spatial Tone Mapping Operator for High Dynamic Range Images. *Proceedings of IS&T/SID's 13th Color Imaging Conference*, CIC 2005.

Foster A, Resnikoff S., 2005 The impact of Vision 2020 on global blindness. *Eye* 2005, 19:1133-1135.

Massof R. W., Rickman D. L., 1992, Obstacles encountered in the development of the low vision enhancement system. Optom Vis Sci, 69:32-41.

Morillas, C., Romero, S., Martínez, A., Pelayo, F. J., Ros, E., Fernández, E. 2007, A Design Framework to Model Retinas. BioSystems 87: 156-163.

NVIDIA Corporation, 2007, Accelerating MATLAB with CUDA using MEX files.

NVIDIA Corporation, 2009. NVIDIA CUDA C Programming Best Practices Guide 2.3

Pelaez-Coca M. D., F. Vargas-Martin, S. Mota, J. Díaz, E. Ros-Vidal, 2009 A versatile optoelectronic aid for low vision patients; Ophtalmic and Phsysiological Optics. Vol. 29, pp. 565-572.

Peli, E., 2001, Vision multiplexing: an engineering approach to vision rehabilitation device development. Optom Vis SCI. 78, 304-315.

Peli E., Luo G., Bowers A., Rensing N., 2007, Applications of augmented-vision head-mounted systems in vision rehabilitation, *Journal of the SID* 15/12.

Ureña, R., Morillas C., Gómez-López J.M., Pelayo F., Cobos J.P., 2010, Plataforma Hw/Sw de aceleración del procesamiento de imágenes. Actas I Simposio de Computación Empotrada (SiCE)